Discretionary Aggregation in Segment Reporting: An Examination Using U.S. Manufacturing Plants*

First Draft: Please check with authors prior to citing this manuscript.

Daniel A. Bens
Eller College of Management, University of Arizona, 1130 East Helen Street, Tucson, AZ 85721

Philip G. Berger
Graduate School of Business, University of Chicago, 5807 S. Woodlawn Ave., Chicago, IL 60637

and

Steven J. Monahan
INSEAD, Boulevard de Constance, CEDEP No. 11, F-7705 Fontainebleau Cedex, 77305, France

March 9, 2007

The research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the Chicago Census Research Data Center (RDC). Lynn Riggs of the Chicago RDC has provided excellent advice and guidance throughout this project. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. This paper has been screened to ensure that no confidential data are revealed. Support for this research at the Chicago RDC from the NSF (awards no. SES-0004335 and ITR-0427889) is gratefully acknowledged. Berger also gratefully acknowledges financial support from the University of Chicago's Graduate School of Business and its Neubauer Family Faculty Fellows program. Bens gratefully acknowledges financial support from the Ratoff Family Fellowship at the University of Arizona.

Discretionary Aggregation in Segment Reporting: An Examination Using U.S. Manufacturing Plants

Abstract:

We investigate the use of discretion in segment reporting by examining Census micro-level data on U.S. manufacturing plants, which we then compare with published financial statement line-of-business segment reports of the firms that own the plants. We conduct our tests by grouping a firm's plants that share the same four-digit SIC code into a "pseudo-segment." Our main findings are that Census pseudo-segments are more likely to be aggregated within a financial reporting line-of-business segment when the agency and proprietary costs of separately reporting the pseudo-segment are higher and when firm and pseudo-segment characteristics allow for more discretion in the application of segment reporting rules. We do not find evidence consistent with the level of managerial entrenchment affecting the segment aggregation decision. Finally, preliminary evidence is consistent with pseudo-segments being more likely to be disclosed as financial statement segments when the firm faces greater litigation risk.

Keywords: manufacturing plants; micro-level data; segment reporting; discretionary disclosure; agency costs; proprietary costs

Data Availability: All data are available from public sources, except for the Census micro-level data.

1. INTRODUCTION

This paper investigates what motivates managers to conceal or reveal information about manufacturing plant performance via the aggregation of these plants into externally reported segments. We examine four possible motives for discretionary nondisclosure (i.e., aggregation) of plant information under SFAS 14, namely, agency costs, proprietary costs, managerial entrenchment and litigation risk. We also control for the characteristics of the manufacturing plants and their firms that affect the level of segment reporting flexibility available under SFAS 14's rules.

Prior empirical studies of segment disclosure choice (e.g., Harris 1998; Botosan and Stanford 2005) have focused primarily on examining the traditional proprietary cost motive offered by the literature to explain nondisclosure in general (e.g., Verrecchia 1983) and aggregation of segments in particular (e.g., Hayes and Lundholm 1996). This theory posits that nondisclosure occurs because disclosure would reveal proprietary information of value to competitors, suppliers, or regulators.

Recent work by Berger and Hann (2007) argues, however, that much of the prior evidence consistent with the proprietary cost hypothesis is also consistent with an alternative "agency cost" hypothesis that posits disclosures are withheld as a result of conflicts of interest between managers and shareholders. The agency cost hypothesis is that segment nondisclosure results from managers attempting to reduce the potential costs to them from segment disclosures that reveal underperformance associated with agency problems. Berger and Hann's findings are consistent with the agency cost hypothesis, but are not generally consistent with the proprietary cost hypothesis.

We follow prior literature in hypothesizing that agency and proprietary costs reduce the extent of discretionary disaggregation of plant-level information in segment disclosures. Prior literature generally suggests that another force we investigate, litigation risk, acts in the opposite direction by encouraging greater discretionary disclosure. Finally, the impact of managerial entrenchment on discretionary disclosure is ambiguous. Entrenched managers may pursue disclosure policies that make it difficult to gauge the extent of their rent extraction. Such managers could, however, be less concerned about whether their rent extraction is observed because outsiders are less able to constrain the actions of more entrenched managers.

Part of our investigation of the agency and proprietary cost motives for nondisclosure follows Berger and Hann (2007) by examining how these motives interact with the level of *abnormal plant profits*, which we define as the plant's gross margin relative to that of its industry. Given the limited set of items that are disclosed in segment footnotes, plant profitability is likely the most valuable piece of information that managers might wish to withhold. Following Berger and Hann (2007), we hypothesize that managers tend to withhold the plants with relatively high (low) abnormal plant profits when the proprietary (agency) cost motive dominates [hereafter, the proprietary (agency) cost motive hypothesis]. In addition to examining how agency and proprietary costs manifest themselves indirectly through the impact of abnormal plant profitability on segment disclosure, we also directly examine the impact on the segment disclosure decision of proxies for each of these forces and for litigation risk and managerial entrenchment.

We test our hypotheses using a sample of 3,086 firm-years from 1984 – 1997 with coverage in both Compustat and the Longitudinal Research Database (LRD), maintained by the Center for Economic Studies at the Bureau of the Census. Our main analyses are conducted using logistic regressions at the "pseudo-segment" level. Because our dependent variable in these analyses is based on a comparison of the four-digit SIC code of an LRD plant versus a Compustat segment, we aggregate all of the firm's LRD plants within the same four-digit SIC code together into one pseudo-segment. The dependent variable in our logitistic regressions is an indicator set equal to one for pseudo-segments with four-digit SIC codes that match either the primary or secondary SIC code of a disclosed segment for that firm ("disclosed" pseudo-segments) and zero otherwise ("hidden" pseudo-segments).

We find that Census pseudo-segments are more likely to be hidden within a financial reporting line-of-business segment when the agency and proprietary costs of separately reporting the pseudo-segment are higher and when firm and pseudo-segment characteristics allow for more discretion in the application of segment reporting rules. We do not find evidence consistent with the level of managerial entrenchment affecting the segment aggregation decision. Finally, preliminary evidence is consistent with pseudo-segments being more likely to be disclosed as financial statement segments when the firm faces greater litigation risk.

Our results make three main contributions. First, the LRD data allow us to observe the firm's operations at a more disaggregated level than what is presented in financial reporting segment footnotes. Past research, most notably Harris (1998), Botosan and Stanford (2005) and Berger and Hann (2007), has examined managers' aggregation decisions by evaluating the amount of segment-level information in publicly disseminated financial statements. A fundamental limitation of using publicly available data is that the researcher can only observe the ex-post aggregate data without having access to the underlying raw source data that managers observe when making the aggregation decision.

For example, Harris (1998) counts the number of SIC codes the firm reports to a *Standard & Poor's* survey, and compares that to the number of segments the firm reports in its published financial statements. The approach cannot capture the *magnitude* of the activity in the SIC code. If the level of operations is small within that SIC code, Harris's proxy for expected segment disclosure is measured with error as it treats the small activity as one that requires an individual segment disclosure. We are able to overcome this drawback because we observe the magnitude of sales for each plant, and thus for each of the firm's SIC codes.

Botosan and Stanford (2005) and Berger and Hann (2007) attempt to overcome some of the limitations of only observing ex-post aggregate data by exploiting the change in segment disclosure rules from SFAS 14 to SFAS 131. Arguing that the change lead to more informative segment disclosures, these papers compare the degree of information aggregation originally reported under the old standard to the restatement of those reports under the new standard. However, these two studies are limited by examining only the restated year(s) of reporting under the old standard and by the fact that the new standard is still subject to considerable managerial discretion. The strict confidentiality rules that protect Census data combined with the much greater level of fineness in plant versus segment data make it much less likely that the plant-level data we use to benchmark segment disaggregation were subject to discretion aimed at misleading financial statement users. Our paper is the first to examine how this confidential internally reported line of business information influences an external reporting disclosure decision.

Second, the LRD data allow us to observe both publicly traded and private firms that compete in the same industry. This provides us with two advantages. We are able to measure some variables more accurately than in prior research as a result of observing all firms competing in the industry – e.g., our measures of the Herfindahl index of industry concentration generally indicate considerably less concentration than the measures used in prior studies because about two-thirds of the sales in our industries are made by private firms. Moreover, we also use this unique data set to develop a proprietary cost proxy as the proportion of industry sales made by private firms. Our intuition is that a public firm competing in an industry with more privately held firms would be more likely to mimic the non-disclosure policies of the private competition. Our paper is the first to consider the impact of private firm competition on the external reporting disclosure decision.

Third, most prior research (e.g., Harris 1998; Ettredge et al. 2002) focuses on examining the proprietary costs of segment disclosure. While Berger and Hann (2007) extend prior work by also incorporating agency costs of disclosure in their analysis, we further extend the segment disclosure literature by investigating proprietary, agency, entrenchment and litigation costs.

The rest of the paper is organized as follows. Section 2 discusses related literature and presents our hypotheses. Section 3 first presents our sample selection and describes our data, then explains our research design. Our empirical results are described in section 4 and section 5 concludes.

2. RELATED LITERATURE AND RESEARCH HYPOTHESES

Several prior studies have explored whether proprietary information costs reduce disclosure of industry segments, with mixed results. Consistent with the proprietary cost hypothesis, Harris (1998) finds that operations in less competitive industries are less likely to be reported as industry segments and Botosan and Stanford (2005) provide evidence that, under SFAS 14, managers hide profitable segments operating in less competitive industries. Inconsistent with the proprietary cost hypothesis, Botosan and Harris (2000) do not find any association between proprietary costs and the decision to voluntarily increase segment disclosure frequency. Berger and Hann (2007) generally fail to find results consistent with the hypothesis that proprietary costs were an important motive for withholding line-of-business segments under SFAS 14. Similar to this line of research, our first hypothesis predicts that managers face proprietary cost motives to withhold segment data.

Managers may also face agency costs of segment disclosure, which arise when segment data provide information that is indicative of unresolved agency problems. Segment data provide information about a company's diversification strategy and its transfers of resources across divisions. Prior research finds evidence consistent with internal capital markets in conglomerates transferring funds across segments in a suboptimal manner (Berger and Ofek 1995; Lamont 1997; Shin and Stulz 1998; Rajan, Servaes and Zingales 2000). Several studies indicate that diversified firms trade at a discount relative to stand-alone firms (Lang and Stulz 1994; Berger and Ofek 1995) and that the diversification discount is associated with measures of agency problems (Dennis, Dennis, and Sarin 1997; Berger and Ofek 1999).[1]

Berger and Hann (2003) and Sanzhar (2003) find that firms moving to more disaggregated segment reporting under the mandated change of SFAS 131 experienced an increased diversification discount, implying that managers concealed information about agency problems under SFAS 14.[2] The same implication arises from Bens and Monahan (2004), who find that greater voluntary segment disaggregation is associated with a smaller diversification discount. Berger and Hann (2007) test this implication and find evidence consistent with the withholding of segment data being motivated by the desire to conceal agency problems. Following Berger and Hann (2007), our second hypothesis therefore predicts that managers face agency cost motives to withhold segment data.

---

[1] A considerable debate is ongoing, however, as to whether the diversification discount results from diversification per se as opposed to self-selection effects (in which firms with poorer prospects are more likely to diversify and/or segments with poorer prospects are more likely to be acquired than to stand alone). The view that diversification per se accounts for the discount is usually linked to the view that diversified firms, on average, have greater unresolved agency conflicts than their pure-play peers. The position that self-selection can explain part or all of the discount sometimes attributes the self-selection mainly to efficiency motives (e.g., Campa and Kedia 2002) and sometimes ascribes the self-selection at least partly to agency problems that exist at the diversified firm prior to its diversifying acquisitions (e.g., Villalonga 2004).

[2] SFAS 14 is FASB Statement No. 14, *Financial Reporting for Segment of a Business Enterprise* (FASB 1976). SFAS 131 is FASB Statement No. 131, *Disclosures about Segments of an Enterprise and Related Information* (FASB 1997).

Whereas the agency cost motive to withhold segment data leads to the prediction that segment disclosure decreases as unresolved agency problems increase, the impact of managerial entrenchment on discretionary disclosure is ambiguous. More entrenched managers may pursue disclosure policies that make it difficult to gauge the extent of their rent extraction. Because such managers have greater influence over firm policies, they are better able to both extract rents and hide their extraction. On the other hand, their greater power may make them less concerned about whether their rent extraction is observed, because outsiders are less able to constrain the actions of more entrenched managers. Thus, the relation between the extent of managerial entrenchment and the degree of segment disclosure is ambiguously signed.

Prior studies of the forces that reduce discretionary segment disclosure have not considered the role of litigation risk. The literature examining the link between litigation risk and discretionary disclosure produces conflicting predictions, which are summarized by Healy and Palepu (2001, pp. 422-23). The threat of shareholder litigation for inadequate or untimely disclosures can encourage firms to increase voluntary disclosure. Conversely, litigation potentially reduces incentives to provide disclosure, particularly of forward-looking information, if managers believe that the legal system penalizes forecasts made in good faith because it cannot effectively distinguish between unexpected forecast errors due to chance and those due to deliberate management bias. We view the extent of discretionary segment disclosure as being much more likely to create litigation concerns related to inadequate disclosure rather than to misleading forecasts. Thus, despite the mixed predictions in the broader literature, the applicable prediction in our setting seems quite clear – higher litigation risk will be associated with less withholding of segment data.[3]

[3] The empirical evidence is basically confined to the impact of forward-looking, preemptive, unaudited disclosures such as earnings warnings on litigation risk and litigation cost. The evidence on the impact of more timely disclosure on litigation risk and cost is mixed. Skinner (1994, 1997) finds that firms with bad earnings news are more likely to be subject to litigation and that such firms are more than twice as likely to pre-disclose as are firms with good earnings news. He also provides some indication that litigation costs are lower for firms that pre-disclose than for firms that do not. Francis et al. (1994), however, provide direct evidence regarding the effect of early disclosure on the likelihood of litigation and find that early disclosure increases the probability of a lawsuit. Finally, Field et al. (2005) argue that the Francis et al. finding could be due to the endogeneity between disclosure and

3. SAMPLE SELECTION, DATA AND METHODOLOGY

3.1 Sample Selection & Data

Our sample selection begins with all annual firm observations during 1984 – 1997 for which the firm is listed on Compustat's Annual Industrial, Research, or Full Coverage files and also covered within the Longitudinal Research Database (LRD), maintained by the Center for Economic Studies at the Bureau of the Census. The LRD is made up of two databases. The first is the *Census of Manufactures* (*CM*), which is conducted every five years (our sample includes the years 1987, 1992 and 1997). The unit of analysis is a manufacturing plant, which the Census refers to as an *establishment*. Extremely small establishments are excluded from the *CM*, although the Census does impute estimates for them based on Internal Revenue Service and Social Security Administration data; all other establishments are required by law to respond truthfully to the *CM* (U. S. Code Title 13, § 224). The second LRD database is the *Annual Survey of Manufactures* (*ASM*), which is conducted in all non-*CM* years (in our sample, 1984-1986, 1988-1991 and 1993-1996). All establishments with more than 250 employees as of the most recent *CM* are included in the *ASM* panel. Smaller establishments are selected with a probability proportional to their size. Once selected, the establishment will remain in the *ASM* for the four years following the *CM*. For a more comprehensive description of the LRD, see McGuckin and Pascoe (1988) or the web site of Center for Economic Studies (http://www.ces.census.gov).

The sample begins in 1984 because that is the first year of Compustat industry segment data that we have access to. The sample ends in 1997 because that is the last year in which firms reported segments under the SFAS 14 rules and also the last year in which the Bureau of the Census classified industries using SIC codes.

We also require our sample firms to have more than one establishment and to have establishments that collectively operate in more than one four-digit SIC code. These requirements ensure that all sample firms have multiple pseudo-segments and thus have some potential for aggregation of pseudo-segments

litigation and, using a simultaneous equations method, they find evidence consistent with disclosure deterring some types of litigation.

into financial reporting segments. Because the LRD covers only U.S. manufacturing establishments, we also require that the firm be domiciled in the U.S. and that its primary SIC code on Compustat fall within the manufacturing sector (i.e., between 2000 and 3999). Even so, the firms sometimes have a portion of their production that occur either outside the U.S. or outside of the manufacturing sector. Moreover, while double-counting of inter-segment sales within a firm is eliminated correctly in Compustat, the attempt to eliminate double-counting of inter-plant sales in the Census data is imperfect.[4] Therefore, the final step in our sample selection process is to eliminate observations where the match between the firm's sales activity as captured by the LRD versus Compustat is poor. We do so by eliminating observations where the ratio of the firm's LRD *total value of shipments* (*TVS*) to Compustat sales is less than 0.75 or greater than one.[5]

Following Harris (1998), our main analyses use a dichotomous dependent variable that measures whether or not a set of plants sharing the same four-digit SIC code gets disclosed as the primary or secondary component of a financial statement segment. Under SFAS 14, enterprises were required to classify line-of-business segment information using the *industry approach*. We therefore group the plants of a firm by aggregating all its establishments that share the same four-digit SIC code in a given year into a "pseudo-segment."

Our full sample includes 17,504 pseudo-segments, comprising 3,086 firm-years. The number of pseudo-segments and firms is not evenly divided across the years 1984 – 1997 because the LRD collects much more extensive information every fifth year. Thus, we have additional data items and many more observations available in 1987, 1992, and 1997.

We are interested only in the line-of-business (LOB) segment-reporting choice for two reasons. First, as a practical matter, the LRD data only cover U.S. manufacturing plants and the extent of geographical diversification in these data is thus limited to within-U.S. variation. Second, and more importantly for

[4] This appears to be largely because inter-plant sales for a plant may be recorded with a value of zero when in fact the plant's survey response about inter-plant sales is missing.

our purposes, it is LOB (and not geographic) diversification that is argued in prior research to be associated with agency problems. Our focus on LOB segment disclosure is also consistent with most prior studies of segment disclosure. Under SFAS 14, industry LOB and geographic segment disclosures are separately reported, and hence it is easy to identify the LOB segments.

3.2 Research Design

3.2.1 Model of Managers' Segment Reporting Choice

We examine managers' segment reporting decisions by estimating the following logit regression at the pseudo-segment level:[6]

$$\begin{aligned} MATCH = \alpha &+ \beta_1\, I\_PROFIT + \beta_2\, TAKEOVER + \beta_3\, PRIVATE \\ &+ \beta_4\, PROFITADJ + \beta_5\, FSIZE + \beta_6\, TRANSIN + \beta_7\, HERF \\ &+ \beta_8\, RELSIZE + \beta_9\, SEGDIVERSITY + \beta_{10}\, CEN\_CMPSTAT + \varepsilon \end{aligned} \qquad (1)$$

The dependent variable, *MATCH*, is a dichotomous variable with the value of one if the pseudo-segment's four-digit SIC code from Census matches a primary or secondary SIC code of a Compustat segment reported by the firm, and zero otherwise. *I_PROFIT*, our measure of abnormal profits, is measured as the pseudo-segment's gross profit margin less the industry average gross profit margin of its four-digit SIC code, with the difference then divided by the standard deviation of gross profit margin in the industry. Gross profit margin is calculated, using LRD fields, as (*TVS* – *CM* – *SW*) / *TVS*, where *TVS* is total value of shipments, *CM* is cost of materials and *SW* is salaries and wages.

We recognize that industry-adjusted gross profit margin captures a pseudo-segment's profit margin rather than total pseudo-segment profitability (i.e., it ignores asset turnover). By industry-adjusting, however, we reduce the likelihood of our gross profit margin measure having a low correlation with ROA because the negative correlation between total asset turnover and gross margin is smaller within industries than across industries. We will test that the intra-industry correlation between gross margin and ROA is

---

[5] *TVS* is the Census' term for the value of establishment sales. It is defined as the "selling value f.o.b. plant after discounts and allowances and excluding freight charges." (*Value of Product Shipments: 2005 – Annual Survey of Manufactures*, Appendix A, issued November 2006).

[6] As will be discussed in the text that follows, model (1) can only be estimated in census years 1987, 1992 and 1997 because of the limited availability of the *TRANSIN* variable. Therefore, we also estimate the augmented version of the model excluding that variable but using our full sample of 1984-1997.

relatively high. We do not attempt to calculate an ROA type measure because the Census stopped collecting asset values in the *ASM* after 1985.

While our measure of abnormal profits is subject to potential measurement error, it has the same advantages over the profit measures used in prior studies as the abnormal profit measure used in Berger and Hann (2007). Specifically, *I_PROFIT* is a direct measure of pseudo-segment profits that captures how well a pseudo-segment performs relative to its industry peers. Other measures of segment profitability such as the speed of profit adjustment measure developed in Harris (1998) generally reflect industry profitability characteristics, which are not hidden. Publicly available industry information is likely of secondary importance in managers' reporting decisions compared with a direct measure of the pseudo-segment's performance. Prior studies (e.g., Botosan and Stanford 2005) have also used firm-level abnormal profits to capture potential proprietary costs. Firm-level profit measures have less measurement error because they are based on GAAP definitions and are not affected by within-firm allocations. On the other hand, it is pseudo-segment profits rather than firm-level profits that managers may try to hide via aggregation. Our measure of pseudo-segment profits is therefore a more relevant measure for studying discretionary disclosure, although this enhanced relevance carries with it the problem of potential measurement error.

Our proxy for the level of managerial entrenchment is the indicator variable, *TAKEOVER*, that equals one (zero) if the firm is, in the year of observation, incorporated in a state with (without) business combination laws that limit an outsider's ability to subject the firm's managers to discipline via the threat of a takeover. Using the Census Bureau's LRD database, Bertrand and Mullainathan (2003) find the adoption of such takeover laws is associated with a significant change in the behavior of firms incorporated in such states. Specifically, management of firms in these states raised wages (especially for white collar workers), avoided the retirement of obsolete assets, and decreased investment in new plants.

PRIVATE is a measure of the potential proprietary information costs of disaggregated disclosure. Our prediction is that a publicly traded firm may be most likely to attempt to mimic the nondisclosure advantage of its private peers by not separately disclosing a line of business when that business unit faces a relatively high level of private firm competition. This variable is measured as the ratio of the sum of *TVS* for all private firms in the four-digit SIC code to the sum of *TVS* for all firms in that SIC code, per the LRD. We classify a Census firm as private if it cannot be linked to a Compustat firm, and we calculate *PRIVATE* as of the most recent census year.[7]

PROFITADJ is an industry abnormal profit adjustment measure constructed based on Harris (1998). It is intended to capture the speed with which abnormal profits are driven down to a normal rate of return. We estimate the persistence of abnormal profits for each industry using the following industry pooled cross-sectional time-series regression over the period 1984 to 1997:

$$X_{ijt} = \beta_{0j} + \beta_{1j}\,(D_n X_{ijt-1}) + \beta_{2j}\,(D_p X_{ijt-1}) + \varepsilon_{ijt} \qquad (2)$$

where,

X_{ijt} The year t difference between plant i's gross margin and the mean gross margin for its four-digit industry, j.
D_n An indicator variable with the value of 1 if X_{ijt-1} is less than or equal to zero; 0 otherwise.
D_p An indicator variable with the value of 1 if X_{ijt-1} is greater than zero; 0 otherwise.

The coefficient estimate of $D_p X_{ijt-1}$ (i.e., β_{2j}) is used to measure the speed of adjustment for positive abnormal ROA in industry *j*, with greater β_{2j} indicating a slower rate of abnormal profit adjustment. We then assign the industry measure to each pseudo-segment by four-digit SIC code to obtain the industry abnormal profit adjustment rate applicable to each pseudo-segment.

Firm size is an important variable in many settings and may be associated with discretionary disclosure policy for a number of reasons. Prior research indicates that greater size is generally associated with a higher level of disclosure (e.g., Lang and Lundholm 1993). A relation between size and discretionary disclosure may be attributable to firm size being a proxy for litigation risk, because larger firms have more assets and thus make more attractive targets for litigation (Kasznik and Lev 1995). We use firm size, *FSIZE*, as a preliminary proxy for litigation risk while recognizing that it may also capture other effects such as economies of scale in financial accounting data collection and dissemination (we

[7] Restricting our sample to U.S. firms is an advantage with respect to studying the impact of private competitors on the segment aggregation decisions of public firms. In many European countries, private companies must make their

will therefore introduce more focused litigation risk proxies in a future draft). We measure *FSIZE* as the natural logarithm of total firm assets (in $ millions) from Compustat.

TRANSIN is an indicator variable that measures whether a pseudo-segment receives transfers from the remainder of the firm. We follow Rajan et al. (2000) in using the difference between the investment a pseudo-segment makes when it is part of a multi-pseudo-segment firm and the investment it would have made had it been on its own as our proxy for transfers made (if negative) or received (if positive). As a first approximation of the investment a pseudo-segment would have made if on its own, we use the *investment ratio* of the remaining LRD plants (other than those in the pseudo-segment being assessed) in the pseudo-segment's four-digit SIC code. We define the investment ratio as the average of the ratio of net new plants (i.e., plants added less plants removed) to end-of-period plants. Because the time-series coverage of individual plants is incomplete in *ASM* years, we measure the investment ratio using the four *CM*s available to us: 1982, 1987, 1992 and 1997. Thus, we measure the investment ratios for 1987, 1992 and 1997 based on the five year change in number of plants.[8]

It is likely, however, that our sample firms have more funds overall than the other firms included in the LRD, because many of the LRD firms not in our sample are small, private companies and thus likely face a higher cost of capital. By measuring transfers as the difference between the investment ratio of a pseudo-segment and the investment ratio of plants in the same industry, we would incorrectly treat these additional funds as a transfer between pseudo-segments rather than as a net addition to all pseudo-segments. To correct for this, we further subtract the industry-adjusted investment ratio averaged across the pseudo-segments of the firm from the pseudo-segment's industry-adjusted investment ratio. Following Rajan et al., we refer to this industry- and firm-adjusted investment ratio as the *adjusted investment ratio*. The adjusted investment ratio proxies for the transfers the pseudo-segment makes (if negative) or receives (if positive). It is computed as:

---

financial statements publicly available and private and public companies are legally subject to the same reporting requirements.

[8] We use number of plants as a measure of investment rather than a more traditional measure such as CAPX because our review of these data in the *ASM* and *CM* suggested that these fields were frequently incomplete.

$$\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} - \sum_{j=1}^{n} w_j \left(\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} \right), \tag{3}$$

where *NP* refers to net new plants added during the five year period, *EP* refers to the number of plants at the end of the five year period, *j* indexes pseudo-segments, *i* indexes industries, *n* is the number of pseudo-segments in the firm and w_j is pseudo-segment *j*'s share of total firm sales. If the computation from expression (3) results in a positive value, the pseudo-segment in question is classified as receiving transfers from the rest of its firm and so *TRANSIN* is set equal to 1. Otherwise, *TRANSIN* is set equal to 0.

Control Variables

We include *HERF*, the Herfindahl index, to control for industry competition. *HERF* is widely used as a measure of industry concentration and competition. It is calculated as follows:

$$HERF = \sum_{i=1}^{n} [s_i / S]^2 \tag{4}$$

Where: s_i = firm *i's TVS* in the four-digit SIC code,
S = the sum of *TVS*, s_i, for all firms in the four-digit SIC code,
n = the number of firms in the industry.

HERF can vary between zero and one. The higher the value of *HERF* is, the higher (lower) the current level of industry concentration (competition) for the industry. We then match the industry measure to each pseudo-segment by four-digit SIC code to obtain the pseudo-segment's industry Herfindahl index.

The next two control variables, *RELSIZE* and *SEGDIVERSITY*, capture the firm's ability to aggregate the pseudo-segment within its financial statement segments. *RELSIZE* is the ratio of pseudo-segment *TVS* to firm *TVS*. If a major goal of segment aggregation is to hide information about abnormal profitability, aggregating small pseudo-segments with larger ones will make it difficult to infer the abnormal profitability of the small pseudo-segment. *SEGDIVERSITY* is a measure of diversity. Under Statement 14, firms with operations in similar industries were afforded greater discretion to aggregate segment information than those with operations in diverse industries. We measure segment diversity as the ratio of the number of unique two-digit SIC codes across pseudo-segments to the total number of pseudo-segments.

Recall from our sample selection that we eliminated all observations where the ratio of the firm's LRD *TVS* to Compustat sales is less than 0.75 or greater than one. Nevertheless, for the observations within our sample we are left with some observations where the LRD *TVS* figure is somewhat less than the Compustat sales number. When these figures do not match, it potentially introduces measurement error because the Compustat segments then include more total sales activity than the Census pseudo-segments. We address this potential measurement error with the variable *CEN_CMPSTAT*, which is calculated as the ratio of total firm *TVS* from Census to total firm sales from Compustat.

3.2.2 Disentangling the Agency and Proprietary Cost Motives

Estimating equation (1) on the entire sample allows us to draw inferences about the extent to which, on average, proprietary or agency cost motives affect segment aggregation decisions. To the extent that one motive or the other dominates, the average effects in the entire sample should reveal this. However, for variables for which a signed prediction arises from only one of the motives (or for which oppositely signed predictions arise from each motive), the role of both motives may be clouded by looking only at average effects within the full sample.

To separately identify agency and proprietary cost motives that may be hidden within the average results for the full sample, it is important to partition our sample such that one motive is likely to dominate the other. Therefore, we follow Berger and Hann (2007) in splitting our observations into two samples: an agency cost (AC) motive sample (where the AC motive dominates) and a proprietary cost (PC) motive sample (where the PC motive dominates). Absent agency problems managers act in the best interests of shareholders and would only choose nondisclose (i.e., aggregation) if the proprietary costs of disclosure outweighed the capital market benefits of disclosure. Thus, our PC motive sample includes all firms where the agency cost motive to withhold pseudo-segment data is relatively weak. When the agency cost motive is strong, managers do not act in the best interests of shareholders and would not be motivated by proprietary cost concerns that conflict with the agency cost motive. Thus, our AC motive sample includes all firms where the agency cost motive to withhold segment information is relatively strong.

To identify a sample of firms where the AC motive likely dominates, we rely on prior literature, which finds evidence of a diversification discount and suboptimal cross-segment transfers. Specifically, in the context of pseudo-segment disclosure, we argue that managers of firms with inefficient cross-pseudo-segment transfers likely face agency cost motives to withhold pseudo-segment data. One approach we take to constructing an AC motive sample is to follow Rajan, Servaes and Zingales (2000) in measuring the diversity of resource-weighted investment opportunities (*DIVERSITY*). It is important to note that Rajan et al. use *DIVERSITY* to capture within-firm agency costs and that they do not consider the firm's (or top managers') incentives to hide these within-firm agency costs from outsiders. In fact, they use SFAS 14 segment data at face value to conduct their empirical tests. We are thus extending the line of reasoning in Rajan et al. by arguing that firm-level managers have incentives to hide these within-firm agency costs from outsiders. To the extent this is true, top management has more desire to aggregate segments when *DIVERSITY* is greater.

Our measure is the standard deviation of the sales-weighted abnormal profitabilities of four-digit SIC codes in which the firm's pseudo-segments operate divided by the equally weighted average abnormal profitability of the four-digit SIC codes in which the firm's pseudo-segments operate:

$$DIVERSITY = \frac{\sqrt{\sum_{j=1}^{n} \frac{(w_j z_j - \overline{wz})^2}{n-1}}}{\frac{\sum_{j=1}^{n} z_j}{n}}, \qquad (5)$$

where w is the sales weighting, z is the abnormal profitability of a four-digit SIC code, j indexes pseudo-segments, n is the number of pseudo-segments in the industry and both w_j and z_j are measured at the end-of-the-period. We measure z for each four-digit SIC code as the industry average gross profit margin divided by the standard deviation of gross profit margin in the industry.

Sales-weighting is used in the numerator to proxy for differences in the resources available to the different pseudo-segments of the firm in the absence of cross-pseudo-segment transfers. Abnormal profitability is a proxy for investment opportunities.[9] Rajan et al. examine diversity of opportunities by developing a model of a two division firm that contains both negative and positive aspects of having more than one division.[10] Consistent with their model's predictions, Rajan et al. find that greater diversity of resource-weighted investment opportunities is associated with: (1) greater transfers from segments with above-average resource-weighted investment opportunities to segments with below-average resource-weighted investment opportunities, (2) decreased efficiency of internal capital market allocations and (3) a greater diversification discount. Therefore, our first approach for partitioning firms into AC and PC motive samples is to classify a firm within the AC motive when the firm's value of *DIVERSITY* is in the top quintile of the distribution for this variable, to classify it within the PC motive when its value of *DIVERSITY* falls within the bottom quintile, and to exclude the firm from consideration when its value of *DIVERSITY* is within the second through fourth quintiles.

While the diversity of investment opportunities has been connected theoretically and empirically to the existence of inefficient funds transfers across business units, it is not actually a measure of the extent of such transfers. We therefore use a second approach for identifying firms with the AC motive for segment aggregation in which we measure the transfers of resources across pseudo-segments, assess the efficiency of the transfers, and classify firms which engage in inefficient transfers as having the AC motive.

The first step in this process is estimating transfers of resources across pseudo-segments. We estimate transfers using the *adjusted investment ratio* of expression (3) that was described earlier in developing our *TRANSIN* variable.

---

[9] We also calculate an alternative version of *Diversity* that, similar to Rajan et al. (2000), uses industry median Tobin's q to proxy for investment opportunities.

[10] In their model, when diversity is low the transfers between the two divisions are value-enhancing, but as diversity increases the transfers become value-decreasing. The transfers in the model occur to prevent a third-best outcome of a "defensive" investment by the division with better resources and opportunities, where a defensive investment offers lower risk-adjusted returns than alternative investments but protects the investing division better against potential expropriation of its economic surplus by the other division. With no transfers between the two divisions, defensive investments would be made by a division when its expected surplus is considerably greater than that of the other division (i.e., when resource-weighted opportunities are considerably greater).

After measuring transfers of funds across pseudo-segments with the adjusted investment ratio, we then use Tobin's q ratios to assess the efficiency of the transfers. We compute q ratios for 1985, 1990 and 1995 for each single-segment Compustat firm in every four-digit SIC code that has at least one Compustat segment in our sample in that year. Our q ratios are computed using the Lindenberg and Ross (1981) methodology and the specific assumptions of Hall et al. (1988). We then assign to each pseudo-segment for each five-year investment window the industry median q ratio of the Compustat single-segment firms that operate in the same industry as the pseudo-segment. Thus, for the 1983 – 1987 investment window we assign the 1985 median q ratio of the pseudo-segment's industry, for 1988 – 1992 the 1990 industry median q ratio, and for 1993 – 1997 the 1995 industry median q ratio. We use the finest industry match that provides at least 5 single-segment Compustat firms with which to calculate the industry median q.

The efficiency of the transfers is then assessed by weighting the transfer to a pseudo-segment by the difference between the pseudo-segment's Tobin's q ratio and the average q ratio in the firm. Under the assumption that the industry median q is a good proxy for the marginal q of a pseudo-segment in that industry, this weighting attributes an incremental market value to each transfer. Following Rajan et al. (2000), we add the weighted transfers across all the pseudo-segments of a firm in a year, and call the sum the *relative value added by allocation* or *RELVALUE*, because it represents a measure of the value consequences of the allocation policy of a diversified firm based on the relative investment opportunity sets of its pseudo-segments. It is given by:

$$\frac{\sum_{j=1}^{n} EP_j(q_j - \bar{q})\left(\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} - \sum_{j=1}^{n} w_j\left(\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j}\right)\right)}{EP}. \qquad (6)$$

The computation of value added by allocation in expression (6) uses the firm's average q ratio to determine whether a pseudo-segment has strong or weak investment opportunities relative to the other pseudo-segments of the firm. Expression (6) multiplies the relative strength of the pseudo-segment's investment opportunities by the measure of transfers obtained from the adjusted investment ratio of expression (3). Because the adjusted investment ratio is designed to avoid overstating transfers between pseudo-segments, it tends to underestimate the value sample firms add via raising capital. For example, the relatively large and publicly traded sample firms can likely raise more external funds than the small, private firms that are generally a substantial portion of each LRD industry. If so, sample firms can thus invest more on average than their industry medians across all pseudo-segments (potentially creating value in all pseudo-segments) and our expression (6) measure does not capture this effect. Therefore, we alternatively compute the value added by allocation by measuring the transfer as the difference between actual pseudo-segment investment and industry peer investments and weighting it by the difference between the pseudo-segment's q ratio and one. This measure, called the *absolute value added by allocation* or *ABSVALUE,* is given by:

$$\frac{\sum_{j=1}^{n} EP_j(q_j - 1)\left(\frac{NP_j}{EP_j} - \frac{NP_j^i}{EP_j^i}\right)}{EP}. \tag{7}$$

4. EMPIRICAL RESULTS

4.1 Analysis of Managers' Segment Reporting Decisions

Descriptive Statistics and Sample Distribution

Table 1 presents descriptive statistics about the sample size and the degree of disaggregation within the Census data and the Compustat data. Our 3,086 firm-year observations represent 962 unique firms. The average number of Census pseudo-segments per firm, at 5.8, is three times the corresponding average of 1.9 Compustat segments per firm and the standard deviation of pseudo-segments per firm is more than five times the standard deviation of segments per firm. Thus, even after grouping plants together by four-digit SIC code, the Census data retain both a higher level of disaggregation than the Compustat segment data and much more variability in the extent of disaggregation.

Table 2 provides descriptive statistics for our main variables. The unit of analysis in this table, and generally for the remaining tables, is the pseudo-segment of the firm, defined as the aggregation of all establishments within a firm that share a common four-digit SIC code. About 37% of the pseudo-segments are "disclosed" as Compustat segments, in that they have a four-digit SIC code that matches either the primary or secondary SIC code of at least one of the firm's Compustat segments for that year (i.e., *MATCH* =1). The Census firms that are also on Compustat are considerably more profitable than their industry peers, with the mean of 0.08 for *I_PROFIT* indicating gross profits of the sample firms exceed those of their industry peers by an average of 0.08 standard deviations. *RELSIZE* averages 0.17, but has a median value of just 0.06, indicating that while pseudo-segments have sales that on average represent about 17% of firm sales, at the median this figure is only about 6%. Thus, the majority of the pseudo-segments represent a small enough portion of the firm to easily be aggregated within a financial statement segment under SFAS 14 rules that require an industry segment to be separately reported if the segment's revenue, earnings, or assets are at least 10 percent of the combined total for that item across all of the firm's industry segments.[11]

The portion of the Compustat firm captured by the LRD is generally less than 100%, as shown by the mean and median value of *CEN_CMPSTAT* being 0.87. The mean value of *TAKEOVER* is 0.66, indicating that about two-thirds of the observations occur when the firm is incorporated in a state with business combination laws that limit the threat of hostile takeover. Finally, the average pseudo-segment operates in a four-digit SIC code in which about 65% of the industry's sales is made by private firms.

[11] More precisely, paragraph 15 of SFAS 14 requires an industry segment to be classified as a reportable segment if it satisfies one or more of the following tests when the tests are applied separately for each fiscal year for which financial statements are presented: (a) Its revenue (including both sales to unaffiliated customers and intersegment sales or transfers) is 10 percent or more of the combined revenue (sales to unaffiliated customers and intersegment sales or transfers) of all of the enterprise's industry segments. (b) The absolute amount of its operating profit or operating loss is 10 percent or more of the greater, in absolute amount, of: (i) The combined operating profit of all industry segments that did not incur an operating loss, or (ii) The combined operating loss of all industry segments that did incur an operating loss. (c) Its identifiable assets are 10 percent or more of the combined identifiable assets of all industry segments. SFAS 14 goes on in paragraph 16 to state that interperiod comparability could require an industry segment to (not) be reportable even if it falls (above) below the 10 percent cut-offs in the currently reported periods. SFAS 14 also allows in paragraph 19 for "the most closely related" reportable segments to be combined into broader reportable segments "to the extent necessary to contain the number of reportable segments within practical limits."

Univariate Analysis

In table 3, we examine how the descriptive statistics for our key explanatory variables differ between the observations for which the dependent variable, *MATCH*, equals one and those for which *MATCH* equals zero. Several striking differences emerge.

The median value of *RELSIZE* is 0.23 for disclosed pseudo-segments versus 0.03 for those not separately disclosed as financial statement segments. This indicates that disclosed pseudo-segments generally contribute more than 10 percent of firm sales, a trigger point for separate segment disclosure under SFAS 14, whereas undisclosed pseudo-segments generally do not. Thus, it is critical to control for *RELSIZE* in our multivariate tests. Nevertheless, the large standard deviation for *RELSIZE* of 0.14 (0.29) among nondisclosed (disclosed) pseudo-segments indicates that a considerable portion of pseudo-segments with *RELSIZE* above (below) 0.10 do not (do) get disclosed as separate segments. Thus, while important, *RELSIZE* is not deterministic with regard to the pseudo-segment disclosure decision.

The distribution of *I_PROFIT* differs in an interesting manner across the undisclosed and disclosed pseudo-segments. The median value of *I_PROFIT* is 0.10 for undisclosed pseudo-segments, whereas it is about 40% larger at 0.14 for the disclosed pseudo-segments. This indicates pseudo-segments are more likely to be aggregated when abnormal profitability is lower, consistent with the AC motive. However, despite having lower median and mean values for *I_PROFIT*, the nondisclosing firms have a larger standard deviation for this variable. The greater variability of abnormal profitability within the nondisclosing sample indicates that there may be a mixture of motives with regard to disclosing pseudo-segments with abnormal profitability.

PROFITADJ has higher mean and median values for the nondisclosing sample. Because higher values of *PROFITADJ* indicate a slower speed of abnormal profit adjustment, this difference indicates that pseudo-segments with a slower speed of abnormal profit adjustment are less likely to be disclosed, consistent with the findings in Harris (1998), Botosan and Stanford (2005), and Berger and Hann (2007). *PRIVATE* also has higher mean and median values for the nondisclosing sample. This indicates that pseudo-segments that are not disclosed are from industries in which a higher portion of industry sales is made by private firms, which have much lower disclosure requirements than public U.S. firms. *HERF* averages just 0.06 for both undisclosed and disclosed pseudo-segments. In addition to indicating that industry concentration does not differ between the undisclosed and disclosed observations, these summary statistics indicate that 4-digit industry concentration measures tend to be extremely low when including the numerous private firms available in the LRD database.

Finally, the mean values of *TRANSIN* indicate that 56% of undisclosed pseudo-segments receive transfers, whereas only 48% of disclosed pseudo-segments are transfer recipients. Thus, the univariate statistics indicate that pseudo-segments that receive transfers are less likely to be disclosed.

Multivariate Logit Regression Analysis

Table 4 presents the results from the multivariate logitistic regression analysis of our base model pooled regression. Because the pooled regression uses the full set of observations during 1984 – 1997, the base model slightly modifies equation (1) by omitting the *TRANSIN* variable, which can only be calculated for the years 1987, 1992 and 1997. The first column of numbers presents the coefficient estimates and the second column the two-tailed p-values[12] The main variables of interest are *I_PROFIT*, *TAKEOVER*, *PRIVATE* and *PROFITADJ*.

Before we turn to the estimates for the main variables, the results for several of the control variables are worth noting. First, the significantly negative estimate on *HERF* is consistent with pseudo-segments being less likely to be disclosed when they are from industries that are more concentrated. We do not have a prediction for the sign of *HERF*, however, and later tables show that the significantly negative estimate here turns out to be fragile. Second, the coefficient estimate on *RELSIZE* is positive and extremely significant, consistent with pseudo-segments that represent a larger portion of their firm's sales being more likely to be disclosed. This is not surprising, as smaller pseudo-segments are probably easier to "hide" as a result of being less visible to investors and competitors. Moreover, SFAS 14's rules result

[12] In a future draft we will also report the corresponding change in probability of the segment being disclosed given a one standard deviation change in the independent variable, calculated as the product of the variable's marginal effect and its standard deviation. Because we include multiple firm observations in our regression analysis, in a future draft we will also present p-values using robust (Huber-White) standard errors that correct for firm clustering.

in auditors being less likely to have concerns about aggregating segments that represent less than 10% of firm sales.

SEGDIVERSITY has a coefficient estimate that is significant at better than the .01 level in the predicted positive direction. This result suggests that the degree of pseudo-segment diversity is an important factor that affects the firm's ability to aggregate pseudo-segments within SFAS 14 financial reporting segments. The finding is perhaps unsurprising, as we expect firms with operations in diverse industries to be less able to aggregate their segment information than firms with operations in similar industries under the SFAS 14 industry-based segment reporting regime.

We draw some comfort from the fact that the estimate on *CEN_CMPSTAT* is not different from zero at conventional significance levels. This indicates that including firm-years where Census *TVS* capture as little as 75% of the total firm sales per Compustat does not distort our inferences about which Census pseudo-segments are disclosed as Compustat segments.

In summary, before turning to our hypothesis tests, we conclude that our dependent variable capturing the internal pseudo-segment disclosure is reasonably associated with firm fundamentals that are likely to be "natural" (i.e., non-strategic) determinants of the disclosure decision.

Turning to our primary variables, the significantly positive coefficient estimate on *FSIZE* indicates that a pseudo-segment is more likely to be separately disclosed when it is part of a larger firm. This is consistent with prior findings that size is associated with a higher level of disclosure (e.g., Lang and Lundholm 1993) and may in part be attributable to the notion that firm size can be viewed as a proxy for litigation risk because larger firms have more assets and thus make more attractive targets for litigation (Kasznik and Lev 1995; Jones and Weingram 1996a; Field et al. 2005). To test the litigation risk hypothesis more thoroughly, we plan to add an additional proxy (or proxies) for litigation risk in future work. Candidates for the new variable(s) include the following: (1) Prior stock returns and return volatility (Jones and Weingram 1996a), although doubt is cast on these variables by the findings in Francis et al. (1994) and Field et al. (2005). (2) An appropriately scaled share turnover (e.g., measured as the probability that a share has been traded at least once during the past year), because lawsuit damages are an increasing function of the number of shares that trade at misleading prices (Jones and Weingram 1996a). (3) An indicator for whether the observation is from the technology sector (in our sample, SIC codes 2833-2836, 3570-3577, and 3600-3674) based on evidence that firms in these industries are sued more often both before the Private Securities Litigation Reform Act of 1995 (Jones and Weingram 1996b) and after it (Grundfest and Perino 1997; Field et al. 2005).

The coefficient estimate for *I_PROFIT* is significantly positive. This result indicates that, on average, pseudo-segments with high abnormal profitability are more likely to be disclosed. The result is consistent with the AC hypothesis that pseudo-segments with low abnormal profitability are more likely to be aggregated. The insignificant estimate on *TAKEOVER* is not consistent with more entrenched managers behaving differently with regard to their pseudo-segment reporting decisions.

The significantly negative coefficient estimate for *PRIVATE* shows that pseudo-segments are less likely to be disclosed when a larger portion of the sales in the pseudo-segment's industry are made by private firms. Because private firms are not required to release their financial statements to the public, a publicly traded firm facing competition from private firms may find itself at a competitive disadvantage attributable to mandated SEC disclosures. The strong, negative relation between *PRIVATE* and pseudo-segment disclosure is consistent with public firms being more likely to aggregate a pseudo-segment (and thus reduce disclosure of it downward toward private firm levels) when that pseudo-segment's competition comes more heavily from private firms, consistent with the PC hypothesis.

Finally, we find a negative and statistically significant coefficient estimate on *PROFITADJ*. The finding on *PROFITADJ*, consistent with Harris (1998), suggests that managers tend to withhold the segments that operate in industries in which the firms with above-average profitability maintain their profitability advantage for longer. This finding has been interpreted by Harris (1998) and Botosan and Stanford (2005) as being consistent with the proprietary cost hypothesis. The basis for this interpretation is the view that the industry is less competitive (and proprietary information is thus more valuable) if the top performing firms in the industry have maintained their advantage for longer. Berger and Hann (2007) point out, however, that an alternative interpretation is possible. In particular, the slow convergence of the top performers toward the industry mean performance can be driven by the weaker firms continuing to experience unresolved agency problems (as opposed to the top firms having a competitive advantage via withholding proprietary information). Ex ante, it is unclear which interpretation dominates. We therefore interpret the negative estimate on *PROFITADJ* in Table 4 as being consistent with the hidden pseudo-segments having been aggregated under SFAS 14 for either agency cost or proprietary cost reasons.

The estimation reported in Table 4 excluded the *TRANSIN* variable from equation (1) because this variable can only be calculated for 1987, 1992 and 1997. Therefore, Table 5 reports estimations of the full equation (1) specification on the set of pooled observations available for 1987, 1992 and 1997. Before turning to the results on the *TRANSIN* variable, it is worth noting that several of the Table 4 results are somewhat sensitive to using only 1987, 1992 and 1997 rather than the entire 1984 – 1997 period. The most sensitive result appears to be that on the control variable *HERF*, which has a significantly positive coefficient estimate in Table 5 in contrast to its significantly negative estimate in Table 4. While retaining the same sign, the estimates on *I_PROFIT*, *PRIVATE*, *SEGDIVERSITY* and *FSIZE* all become much less significant. Note that an untabulated version of Table 5 that excludes *TRANSIN* establishes that all of the differences in results between Tables 4 and 5 are due to using only 1987, 1992 and 1997 in Table 5 rather than to the addition of the *TRANSIN* variable.

With regard to *TRANSIN*, the significantly negative coefficient estimate is consistent with pseudo-segments being less likely to be disclosed when they are receiving transfers. *TRANSIN* currently does not distinguish between transfers that are efficient and inefficient, so this result cannot unambiguously be viewed as more consistent with the AC hypothesis than the PC hypothesis. However, prior research suggests that a substantial portion of the transfers captured by *TRANSIN* are inefficient. We plan to investigate this issue further by assessing the efficiency of the transfers.

The results so far attempt to discern the extent to which the AC and PC motives explain pseudo-segment aggregation, on average, within the sample. This approach may cloud our ability to distinguish between these competing hypotheses because for many of our inferential variables only one of the hypotheses offers a signed prediction and when both hypotheses offer a prediction for a variable the predictions tend to have opposite signs (e.g. *I_PROFIT*). If the full sample includes some firms motivated by agency cost considerations and others motivated by proprietary cost concerns, the effect of each motive will be obscured in a full sample that includes observations with the other motive. Therefore, we conduct a number of tests in which we first create samples in which the AC motive or PC motive is likely to dominate and then perform estimations of models similar to equation (1) separately on both samples.

The first variable we use to group observations into AC and PC motive samples is *DIVERSITY*, calculated as in equation (5). Based on Rajan et al.'s (2000) evidence that higher values of *DIVERSITY* are associated with greater levels of value-decreasing shifts of funds between the firm's segments, we argue that the AC motive for pseudo-segment aggregation is most likely to be present when the value of *DIVERSITY* is relatively high. We therefore classify firms in the highest quintile of the distribution of *DIVERSITY* as having the AC motive for aggregation, and those in the lowest quintile of the distribution are classified as having the PC motive.

Table 6 Panel A presents the results of re-estimating the Table 4 regression for the AC motive sample, whereas Panel B does the same for the PC motive sample. The results are generally consistent with our predictions except that, contrary to our predictions, there is no evidence that high abnormal profit pseudo-segments are more likely to be disclosed within the AC motive sample. The coefficient estimate of 0.05 on *I_PROFIT* in the AC motive sample regression is neither significantly different from zero nor different from the 0.05 estimate obtained for this variable within the PC motive sample.

Several of the results are consistent with the proprietary cost hypothesis within the PC motive sample, but not within the AC motive sample. The coefficient estimate on *PRIVATE* is negative and highly significant within the PC motive sample, whereas the corresponding estimate within the AC motive sample is much less negative and not significantly different from zero. Similarly, to the extent *PROFITADJ* captures proprietary costs rather than agency costs, the more negative and significant estimate on this variable within the PC motive sample is consistent with the PC motive for aggregation being stronger within this sample.

In Table 7 we add *TRANSIN* to the explanatory variables. Panels A and B report estimations of the complete equation (1) specification for the AC motive and PC motive samples, respectively. As in Table 5, these estimations are performed using only observations from 1987, 1992 and 1997 in order to allow *TRANSIN* to be calculated.[13] The results are generally consistent with our predictions. We do not find statistically significant evidence consistent with high abnormal profit pseudo-segments being more likely to be disclosed within the AC motive sample. Nevertheless, the coefficient estimate of 0.12 for the AC motive sample in Panel A is large relative to our other estimations, including that of 0.04 for the PC motive sample in Panel B, even though the Panel A estimate's p-value of 0.22 indicates low statistical significance.

TRANSIN has a significantly negative coefficient estimate in Panel A and an insignificant estimate in Panel B, consistent with our predictions. These results indicate that pseudo-segments receiving transfers are only less likely to be disclosed within the sample where the AC motive dominates. Finally, *PRIVATE* and *PROFITADJ* have significantly negative coefficients in Panel B, but insignificant estimates in Panel A, consistent with our expectations. These results indicate that withholding pseudo-segments for proprietary cost reasons occurs within the PC motive sample, but not the AC motive sample.

Tables 6 and 7 group firms into AC and PC motive samples based on whether *DIVERSITY* is in the top or bottom quintile of its distribution. Although Rajan et al. (2000) find that higher values of *DIVERSITY* are associated with greater levels of inefficient funds transfers between the firm's segments, *DIVERSITY* does not directly measure the efficiency of the firm's resource allocation across pseudo-segments. Therefore, in Table 8 we partition the observations into AC and PC motive samples using *RELVALUE*, the relative value added by allocation. Using this alternate partitioning scheme matters, as some inferences change compared to those from Table 7.

One finding that remains stable is that we find at best marginally significant evidence consistent with high abnormal profit pseudo-segments being more likely to be disclosed within the AC motive sample.

[13] Note that untabulated versions of Tables 8 and 9 that exclude *TRANSIN* establish that all of the differences in results between Tables 6 and 7 versus Tables 8 and 9 are due to using only 1987, 1992 and 1997 in Tables 8 and 9

The evidence is suggestive of this effect, however, with the coefficient estimate of 0.09 for the AC motive sample in Panel A large relative to our other estimations, including that of 0.03 for the PC motive sample in Panel B, even though the Panel A estimate's two-tailed p-value of 0.15 indicates low statistical significance.

TRANSIN has a significantly negative coefficient estimate in both Tables 10 and 11. These findings indicate that firms with both the AC and PC motives for aggregation are more likely to aggregate pseudo-segments that receive funds transfers. The finding for the AC motive sample is consistent with our predictions to the extent that the funds transfers are often inefficient. We did not predict that firms with the PC motive would be more likely to aggregate pseudo-segments receiving transfers – it is not obvious why the PC motive would make such pseudo-segments more likely to be aggregated unless the extent of access to the internal capital market is valuable proprietary information.

PRIVATE has a significantly negative coefficient estimate in Panel B, but an insignificant estimate in Panel A. These results indicate that withholding pseudo-segments for proprietary cost reasons occurs within the PC motive sample, but not the AC motive sample, consistent with our predictions.

PROFITADJ is significantly negative in Panel A but insignificant in Panel B. The negative finding on *PROFITADJ* suggests that managers tend to withhold the segments that operate in industries in which the firms with above-average profitability maintain their profitability advantage for longer. Finding this result only for the AC motive sample is consistent with the slow convergence of the top performers toward the industry mean performance being driven by the weaker firms continuing to experience unresolved agency problems which they wish to obscure via lower disclosure (as opposed to the top firms having a competitive advantage via withholding proprietary information). The inference that *PROFITADJ* reflects the AC motive more than the PC motive is fragile, as the results using the alternate partitioning variable for the AC motive in Table 7 were consistent with *PROFITADJ* reflecting the PC motive more than AC motive. Overall, our evidence is consistent with prior findings by Harris (1998), Botosan and Stanford (2005), and Berger and Hann (2007) that shows *PROFITADJ* is significantly

---

rather than to the addition of the *TRANSIN* variable.

associated with the segment aggregation decision. Our evidence does not, however, clearly disentangle whether this association is driven by the agency cost or proprietary cost motive for segment aggregation.

4.2 Sensitivity Test

The Table 8 Panel A results are all insensitive to partitioning firms into the AC motive using observations where *the absolute value added by allocation* (or *ABSVALUE*) is negative instead of the set where *RELVALUE* is negative. For Panel B, the analogous sensitivity test for partitioning firms into the PC motive results in the following inference changes: *HERF* no longer has a significantly positive coefficient, *PRIVATE* no longer has a significantly negative coefficient, and *TRANSIN* no longer has a significantly negative coefficient. Overall, partitioning the AC vs. PC motive on the basis of *ABSVALUE* rather than *RELVALUE* does not affect the inferences for the AC motive partition, but does alter the inferences for the PC motive group.

5. CONCLUSION

We use a unique data set to present novel findings on an important discretionary financial reporting decision. Specifically, we use confidential U.S. Census Bureau data at the manufacturing establishment level, and examine how these data are aggregated into externally reported financial segments. Our primary unit of analysis is a pseudo-segment, which reflects our aggregation of all firm establishments that share a common four digit SIC code into one unit. We then examine the determinants of whether than pseudo-segment is disclosed as a line of business segment in the parent firm’s external reports, or whether it is aggregated into a non-matching segment.

Our results suggest that at least three forces influence this disclosure decision, after controlling for economic fundamentals of the firm and pseudo-segment. These forces include proprietary costs, agency costs and litigation risks. Ours is the first paper to model this disclosure decision so extensively, and is made possible by our access to Census’ records of firm-reported internal data. While some of our proxies may be impossible for other researchers to measure with external data (e.g., plant-level profitability), others, such as our measure of competition from privately held firms, might be accessible to other researchers (e.g., through Census industry level reports). Future research might be aimed at refining such proxies with external data to assess whether our model generalizes to data publicly available to all researchers.

REFERENCES

Bens, D. and S. Monahan. 2004. Disclosure Quality and the Excess Value of Diversification. *Journal of Accounting Research* 42, 691-730.

Berger, P. and R. Hann. 2003. The Impact of SFAS No. 131 on Information and Monitoring. *Journal of Accounting Research* 41, 163-223.

_____________. 2007. Segment Profitability and the Proprietary and Agency Costs of Disclosure. *The Accounting Review* 82: forthcoming.

Berger, P. and E. Ofek. 1995. Diversification's Effect on Firm Value. *Journal of Financial Economics* 37, 39-65.

_____________. 1999. Causes and Effects of Corporate Refocusing Programs. *Review of Financial Studies* 12, 311-345.

Bertrand, M. and S. Mullainathan. 2003. Enjoying the Quiet Life? Corporate Governance and Managerial Preferences. *Journal of Political Economy* 111, 1043-1075.

Botosan, C. and M. Harris. 2000. Motivations for a Change in Disclosure Frequency and Its Consequences: An Examination of Voluntary Quarterly Segment Disclosures. *Journal of Accounting Research* 38, 329-353.

Botosan, C. and M. Stanford. 2005. Managers' Motives to Withhold Segment Disclosures and the Effect of SFAS No. 131 on Analysts' Information Environment. *The Accounting Review* 80: 751 – 771.

Campa, J. and S. Kedia. 2002. Explaining the Diversification Discount. *Journal of Finance* 57, 1731-1762.

Denis, D., D. Denis, and A. Sarin. 1997. Agency Problems, Equity Ownership, and Corporate Diversification. *The Journal of Finance* 52, 135-160.

Ettredge, M., S. Kwon, and D. Smith. 2002. Competitive Harm and Managers' Attitude Towards SFAS No. 131. *Journal of Accounting, Auditing, and Finance* 17, 93-109.

Field, L., M. Lowry, and S. Shu. 2005. Does Disclosure Deter or Trigger Litigation? *Journal of Accounting and Economics* 39, 487-507.

Financial Accounting Standards Board (FASB). 1976. *Financial Reporting for Segments of a Business Enterprise.* Statement of Financial Accounting Standards No. 14. Norwalk CT (FASB).

Financial Accounting Standards Board (FASB). 1997. *Disclosures about Segments of an Enterprise and Related Information*. Statement of Financial Accounting Standards No. 131. Norwalk CT (FASB).

Francis, J., D. Philbrick and K. Schipper. 1994. Shareholder Litigation and Corporate Disclosures. *Journal of Accounting Research* 32, 137-164.

Grundfest, J. and M. Perino. 1997. *Ten Things we Know and Ten Things we Don't Know about the Private Securities Litigation Reform Act of 1995*. Joint Written Testimony of Joseph A. Grundfest and

Michael A. Perino Before the Subcommittee on Securities of the Committee on Banking, Housing and Urban Affairs, United States Senate on July 24, 1997.

Hall, B., C. Cummins, E. Laderman and J. Mundy. 1988. The R&D Master File Documentation. National Bureau of Economic Research, Boston, MA.

Harris, M. 1998. The Association between Competition and Managers' Business Segment Reporting Decisions. *Journal of Accounting Research* 36, 111-128.

Hayes, R. and R. Lundholm. 1996. Segment Reporting to the Capital Market in the Presence of a Competitor. *Journal of Accounting Research* 34, 261-279.

Healy, P. and K. Palepu. 2001. Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature. *Journal of Accounting & Economics* 31, 405-440.

Jones, C. and S. Weingram. 1996a. The Determinants of 10b-5 Litigation Risk. Stanford Law School Working Paper.

_____________. 1996b. Why 10b-5 risk is higher for technology and financial service firms. Stanford Law School Working Paper.

Kasznik, R. and B. Lev. 1995. To Warn or not to Warn: Management Disclosures in the Face of an Earnings Surprise. *The Accounting Review* 70: 113 – 134.

Lamont, O. 1997. Cash Flow and Investment: Evidence from Internal Capital Markets. *Journal of Finance* 52, 83-109.

Lang, M. and R. Lundholm. 1993. Cross-sectional Determinants of Analyst Ratings of Corporate Disclosures. *Journal of Accounting Research* 31, 246-271.

Lang, L. and R. Stulz. 1994. Tobin's Q, Corporate Diversification and Firm Performance. *Journal of Political Economy* 102, 1248-1280.

Lindenberg, E. and S. Ross. 1981. Tobin's q and industrial organization. *Journal of Business* 54, 1-32.

McGuckin, R. and G. Pascoe. 1988. The Longitudinal Research Database: Status and Research Possibilities. *Survey of Current Business* 68, 30-37.

Rajan, R,, H. Servaes and L. Zingales. 2000. The Cost of Diversity: The Diversification Discount and Inefficient Investment. *Journal of Finance* 55, 35-80.

Sanzhar, S. 2003. Two Puzzles about the Diversification Discount: SFAS #131 and 'Pseudo-Conglomerates.' Working paper, London Business School.

Shin, H. and R. Stulz. 1998. Are Internal Capital Markets Efficient? *Quarterly Journal of Economics* 113, 531-552.

Skinner, D., 1994. Why Firms Voluntarily Disclose Bad News. *Journal of Accounting Research* 32, 38-60.

_____________. 1997. Earnings Disclosure and Stockholder Lawsuits. *Journal of Accounting and Economics* 23, 249-282.

Verrecchia, R. 1983. Discretionary Disclosure. *Journal of Accounting and Economics* 5, 365-380.

Villalonga, B. 2004. Does Diversification Cause the "Diversification Discount"? *Financial Management* 33, 5-27.

Table 1
Firm Level Descriptive Statistics of Number of Segments

Pseudo-Segments			Compustat Segments		
Mean	Median	Standard Deviation	Mean	Median	Standard Deviation
5.8	3.0	6.2	1.9	2.0	1.1

Statistics are from 3,086 firm-years (962 unique firms) spanning 1984-1997. A firm is included if its primary SIC code is in the manufacturing sector (SIC 2000-3999) and it can be matched from Compustat to the Longitudinal Research Database (LRD) of the U.S. Census Bureau. A *pseudo-segment* is defined as all LRD plants for a firm that operate in a common four digit SIC code. *Compustat segments* reflect line of business segments reported by the firm in its external reports filed with the SEC. A firm must have a minimum of two pseudo-segments to be included in our sample, and the ratio of aggregate *total value of shipments* (*TVS*) per the LRD to sales per Compustat must be greater than 0.75 but less than 1.

Table 2
Pseudo-Segment Level Descriptive Statistics

	Mean	Median	Standard Deviation
MATCH	0.37	0.00	0.48
I_PROFIT	0.08	0.12	0.77
TAKEOVER	0.66	1.00	0.47
PRIVATE	0.65	0.68	0.18
PROFITADJ	0.33	0.34	0.37
FSIZE	6.53	6.67	1.56
HERF	0.06	0.04	0.11
RELSIZE	0.17	0.06	0.24
SEGDIVERSITY	0.50	0.46	0.22
CEN_CMPSTAT	0.87	0.87	0.07

Statistics are from 17,504 pseudo-segments of 3,086 firm-years (962 unique firms) spanning 1984-1997 (see Table 1 for additional sample selection criteria). A *pseudo-segment* is defined as all LRD plants for a firm that operate in a common four-digit SIC code. *MATCH* equals one if the pseudo-segment four-digit SIC code matches the primary or secondary SIC code of a Compustat segment, zero otherwise. *I_PROFIT* equals the difference between the pseudo-segment's gross margin and the industry average gross margin, divided by the standard deviation of gross margin across the industry; gross margin is obtained from the LRD database, and equals total value of shipments (*TVS*) less cost of materials less salaries and wages, with the difference scaled by *TVS*. *TAKEOVER* equals one if the firm is headquartered in a state with anti-takeover protection laws, and zero otherwise (classification based on the list of anti-takeover laws in Bertrand and Mullainathan 2003). *PRIVATE* equals the sum of *TVS* across all firms in the pseudo-segment's industry that cannot be linked to Compustat (i.e., we assume they are thus privately held), divided by the sum of *TVS* across all firms in the industry; the variable is only calculated in census years (1982, 1987, 1992 and 1997) and then applied for all sample years until it is re-measured at the next census. In a regression estimated from all firms in the pseudo-segment's industry where the dependent variable is current year gross margin and the two independent variables are lagged gross margin if negative, and lagged gross margin if positive, the coefficient from the positive gross margin realizations is defined as *PROFITADJ*. *FSIZE* equals the natural logarithm of total firm assets in $ millions per Compustat. *HERF* equals the Herfindahl index for the pseudo-segment's industry, defined as the square of each pseudo-segment's *TVS* deflated by total industry *TVS* with the squared terms then summed across all pseudo-segments within the industry. *RELSIZE* equals the *TVS* of the pseudo-segment deflated by the total *TVS* summed across all of the firm's pseudo-segments. *SEGDIVERSITY* equals number of unique two-digit SIC codes across the firm's pseudo-segments deflated by the total number of pseudo-segments. *CEN_CMPSTAT* equals total *TVS* summed across all of the firm's pseudo-segments deflated by total sales of the firm per Compustat.

Table 3
Descriptive Statistics for *MATCH* vs. non-*MATCH* Pseudo-Segments

	Means		Medians		Standard Deviations	
	MATCH = 1	*MATCH = 0*	*MATCH = 1*	*MATCH = 0*	*MATCH = 1*	*MATCH = 0*
I_PROFIT	0.15	0.04	0.14	0.10	0.64	0.83
HERF	0.06	0.06	0.04	0.03	0.06	0.13
RELSIZE	0.33	0.08	0.23	0.03	0.29	0.14
PRIVATE	0.62	0.66	0.65	0.68	0.19	0.18
PROFITADJ	0.31	0.35	0.30	0.34	0.39	0.36
NEWPLANT	-0.04	0.11	-0.01	0.05	1.58	0.81
TRANSIN	0.48	0.56	0.00	1.00	0.50	0.50

Statistics are from 17,504 pseudo-segments of 3,086 firm-years (962 unique firms) spanning 1984-1997 (see Table 1 for additional sample selection criteria), with the exception of the *NEWPLANT* and *TRANSIN* variables that are measured in census years only (1987, 1992 and 1997) for 5,006 pseudo-segments. The table presents descriptive statistics for those pseudo-segments where the four-digit SIC code can be matched to a Compustat segment's primary or secondary industry code (*MATCH* = 1) vs. pseudo-segments that cannot be matched (*MATCH* = 0). All variables are defined in Table 2 except for *NEWPLANT* and *TRANSIN*. *NEWPLANT* equals the proportion of a pseudo-segment's establishments that were not owned by the firm in the previous census year less the proportion of new plants for the entire industry (excluding the firm) less the weighted (by *TVS*) firm average of this difference across all of its pseudo-segments. In equation form:

$$NEWPLANT_j = \frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} - \sum_{j=1}^{n} w_j \left(\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} \right)$$

In the equation, *NP* denotes new plants, *EP* ending plants, *w RELSIZE* for the pseudo-segment, *j* indexes pseudo-segment and *i* indexes the industry. *TRANSIN* equals one if *NEWPLANT* is positive for the pseudo-segment, zero otherwise.

Table 4
Logistic Regression Estimates for Pseudo-Segment Disclosure Decision Excluding *TRANSIN*

	Coefficient	P-value
Intercept	-1.473	0.000
I_PROFIT	0.077	0.002
TAKEOVER	0.006	0.875
PRIVATE	-0.699	0.000
PROFITADJ	-0.237	0.000
FSIZE	0.067	0.000
HERF	-0.631	0.001
RELSIZE	5.543	0.000
SEGDIVERSITY	0.364	0.000
CEN_CMPSTAT	-0.094	0.712

The regression is estimated using 17,504 pseudo-segments of 3,086 firm-years (962 unique firms) spanning 1984-1997 (see Table 1 for additional sample selection criteria). The binary dependent variable is *MATCH* and is defined along with all of the explanatory variables in Table 2. All p-values are two-tailed.

Table 5
Logistic Regression Estimates for Pseudo-Segment Disclosure Decision Including *TRANSIN*

	Coefficient	P-value
Intercept	-1.130	0.038
I_PROFIT	0.060	0.167
TAKEOVER	-0.065	0.377
PRIVATE	-0.301	0.125
PROFITADJ	-0.202	0.034
FSIZE	0.020	0.452
TRANSIN	-0.223	0.001
HERF	1.849	0.003
RELSIZE	5.243	0.000
SEGDIVERSITY	0.279	0.132
CEN_CMPSTAT	-0.328	0.506

The regression is estimated using 5,006 pseudo-segments from Census years 1987, 1992 and 1997 only (see Table 1 for additional sample selection criteria). The binary dependent variable is *MATCH* and is defined along with all of the explanatory variables in Table 2 except for *TRANSIN* which is defined in Table 3. All p-values are two-tailed.

Table 6
Logistic Regression Estimates for Pseudo-Segment Disclosure Decision Excluding *TRANSIN* – Samples Separated by Diversity of Investment Opportunity

Panel A – Firm Year in the High Quintile Diversity: Agency Cost Hypothesis Dominant

	Coefficient	P-value
Intercept	-1.555	0.006
I_PROFIT	0.048	0.393
TAKEOVER	-0.030	0.710
PRIVATE	-0.186	0.356
PROFITADJ	-0.101	0.324
FSIZE	0.135	0.000
HERF	-1.323	0.001
RELSIZE	4.101	0.000
SEGDIVERSITY	0.392	0.024
CEN_CMPSTAT	-0.375	0.491

Panel B – Firm Year in the Low Quintile Diversity: Proprietary Cost Hypothesis Dominant

	Coefficient	P-value
Intercept	-0.958	0.179
I_PROFIT	0.054	0.329
TAKEOVER	0.304	0.003
PRIVATE	-1.517	0.000
PROFITADJ	-0.259	0.033
FSIZE	0.023	0.544
HERF	0.402	0.438
RELSIZE	5.510	0.000
SEGDIVERSITY	0.647	0.009
CEN_CMPSTAT	-0.396	0.533

(continued on next page)

The regression is estimated on two separate sub-samples over the period 1984-1997 (see Table 1 for additional sample selection criteria). The binary dependent variable is *MATCH* and is defined along with all of the explanatory variables in Table 2. Panel A includes 3,513 pseudo-segment observations in the top quintile of a diversity of investment opportunity measure calculated as follows:

$$DIVERSITY = \frac{\sqrt{\sum_{j=1}^{n} \frac{(w_j z_j - \overline{wz})^2}{n-1}}}{\frac{\sum_{j=1}^{n} z_j}{n}}$$

In the equation above, w_j denotes *RELSIZE* for a pseudo-segment, z_j equals the average gross profit margin for the pseudo-segment's industry, and n indexes all pseudo-segments at a given firm. Panel B includes 3,497 pseudo-segment observations in the bottom quintile of *DIVERSITY*. All p-values are two-tailed.

Table 7
Logistic Regression Estimates for Pseudo-Segment Disclosure Decision Including *TRANSIN* – Samples Separated by Diversity of Investment Opportunity

Panel A – Firm Year in the High Quintile Diversity: Agency Cost Hypothesis Dominant

	Coefficient	P-value
Intercept	-1.495	0.159
I_PROFIT	0.123	0.216
TAKEOVER	-0.206	0.219
PRIVATE	0.264	0.488
PROFITADJ	0.042	0.825
FSIZE	0.108	0.027
TRANSIN	-0.346	0.010
HERF	2.627	0.036
RELSIZE	3.938	0.000
SEGDIVERSITY	-0.043	0.894
CEN_CMPSTAT	-0.202	0.841

Panel B – Firm Year in the Low Quintile Diversity: Proprietary Cost Hypothesis Dominant

	Coefficient	P-value
Intercept	-1.348	0.339
I_PROFIT	0.042	0.675
TAKEOVER	0.138	0.445
PRIVATE	-0.882	0.075
PROFITADJ	-0.433	0.065
FSIZE	-0.019	0.775
TRANSIN	0.137	0.450
HERF	1.808	0.296
RELSIZE	5.512	0.000
SEGDIVERSITY	0.595	0.192
CEN_CMPSTAT	-0.037	0.978

(continued on next page)

The regression is estimated on two separate sub-samples in census years 1987, 1992 and 1997 only (see Table 1 for additional sample selection criteria). The binary dependent variable is *MATCH* and is defined along with all of the explanatory variables in Table 2, with the exception of *TRANSIN* which is defined in Table 3. Panel A includes 1,190 pseudo-segment observations in the top quintile of a diversity of investment opportunity measure described in Table 6. Panel B includes 1,065 pseudo-segment observations in the bottom quintile of the diversity measure. All p-values are two-tailed.

Table 8
Logistic Regression Estimates for Pseudo-Segment Disclosure Decision Including *TRANSIN* – Samples Separated by Relative Investment Efficiency

Panel A – Firm Year with Relative Investment Efficiency < 0: Agency Cost Hypothesis Dominant

	Coefficient	P-value
Intercept	-0.873	0.266
I_PROFIT	0.088	0.152
TAKEOVER	-0.042	0.683
PRIVATE	-0.078	0.778
PROFITADJ	-0.336	0.013
FSIZE	-0.012	0.756
TRANSIN	-0.217	0.022
HERF	1.809	0.038
RELSIZE	4.785	0.000
SEGDIVERSITY	0.489	0.063
CEN_CMPSTAT	-0.592	0.402

Panel B – Firm Year with Relative Investment Efficiency >= 0: Proprietary Cost Hypothesis Dominant

	Coefficient	P-value
Intercept	-1.206	0.115
I_PROFIT	0.029	0.640
TAKEOVER	-0.118	0.268
PRIVATE	-0.547	0.054
PROFITADJ	-0.084	0.533
FSIZE	0.053	0.147
TRANSIN	-0.233	0.019
HERF	1.870	0.035
RELSIZE	5.683	0.000
SEGDIVERSITY	0.039	0.882
CEN_CMPSTAT	-0.207	0.764

(continued on next page)

The regression is estimated on two separate sub-samples in census years 1987, 1992 and 1997 only (see Table 1 for additional sample selection criteria). The binary dependent variable is *MATCH* and is defined along with all of the explanatory variables in Table 2, with the exception of *TRANSIN* which is defined in Table 3. Panel A includes 2,556 pseudo-segment observations where relative investment efficiency is estimated as being negative. Panel B includes 2,450 pseudo-segment observations where relative investment efficiency is estimated as being positive. All p-values are two-tailed. In calculating relative investment efficiency, we first calculate the level of new plant investment for each pseudo segment *j* of the firm as follows:

$$NEWPLANT_j = \frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} - \sum_{j=1}^{n} w_j \left(\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} \right)$$

NEWPLANT equals the proportion of a pseudo-segment's establishments that were not owned by the firm in the previous census year less the proportion of new plants for the entire industry (excluding the firm) less the weighted (by *TVS*) firm average of this difference across all of its pseudo-segments. *NP* denotes new plants, *EP* ending plants, *w RELSIZE* for the pseudo-segment, *j* indexes pseudo-segment and *i* indexes the industry. We then calculate the median Tobin's q ratio for each industry using Compustat data as of the middle of the investment period (i.e., 1985, 1990 and 1995). We then calculate *RELVALUE* as follows:

$$RELVALUE = \frac{\sum_{j=1}^{n} EP_j (q_j - \bar{q}) \left(\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} - \sum_{j=1}^{n} w_j \left(\frac{NP_j}{EP_j} - \frac{NP^i_j}{EP^i_j} \right) \right)}{EP}$$

If *RELVALUE* is less than zero then we consider new investments to be inefficient (Panel A observations); otherwise, they are considered to be efficient (Panel B observation).